

Neural Style Transfer

Minju Kim

2022.01.29.SAT

What is Neural Style Transfer?

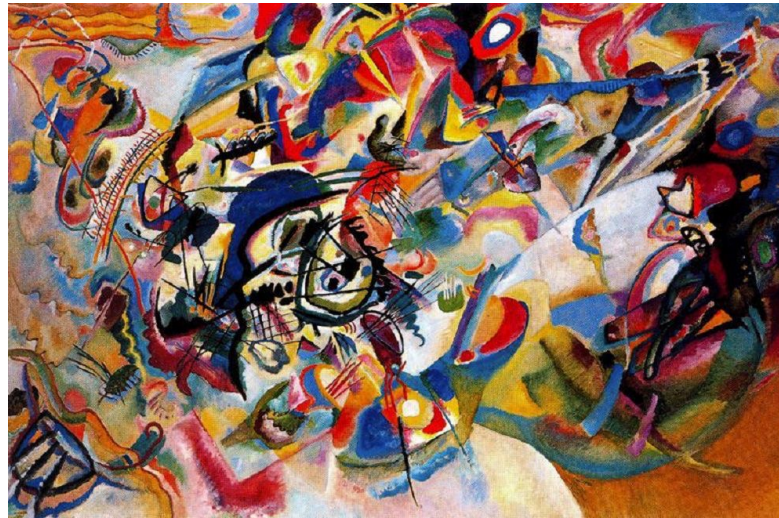
- Neural Style Transfer is the technique of blending style from one image into another image keeping its content intact.
- The only change is the style configurations of the image to give an artistic touch to your image.

Content Image



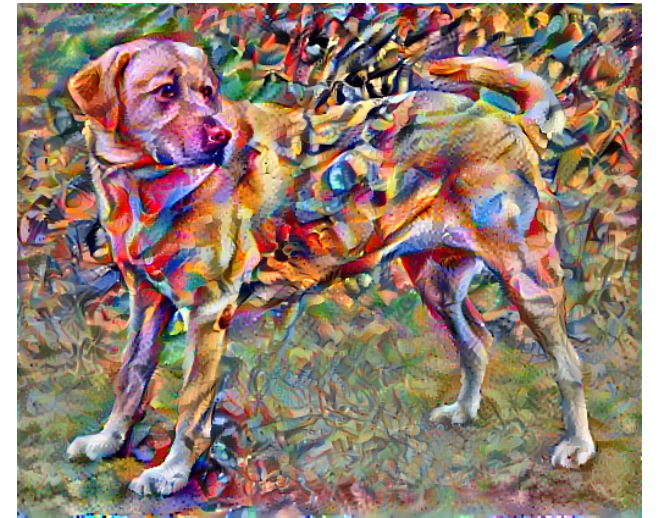
+

Style Image



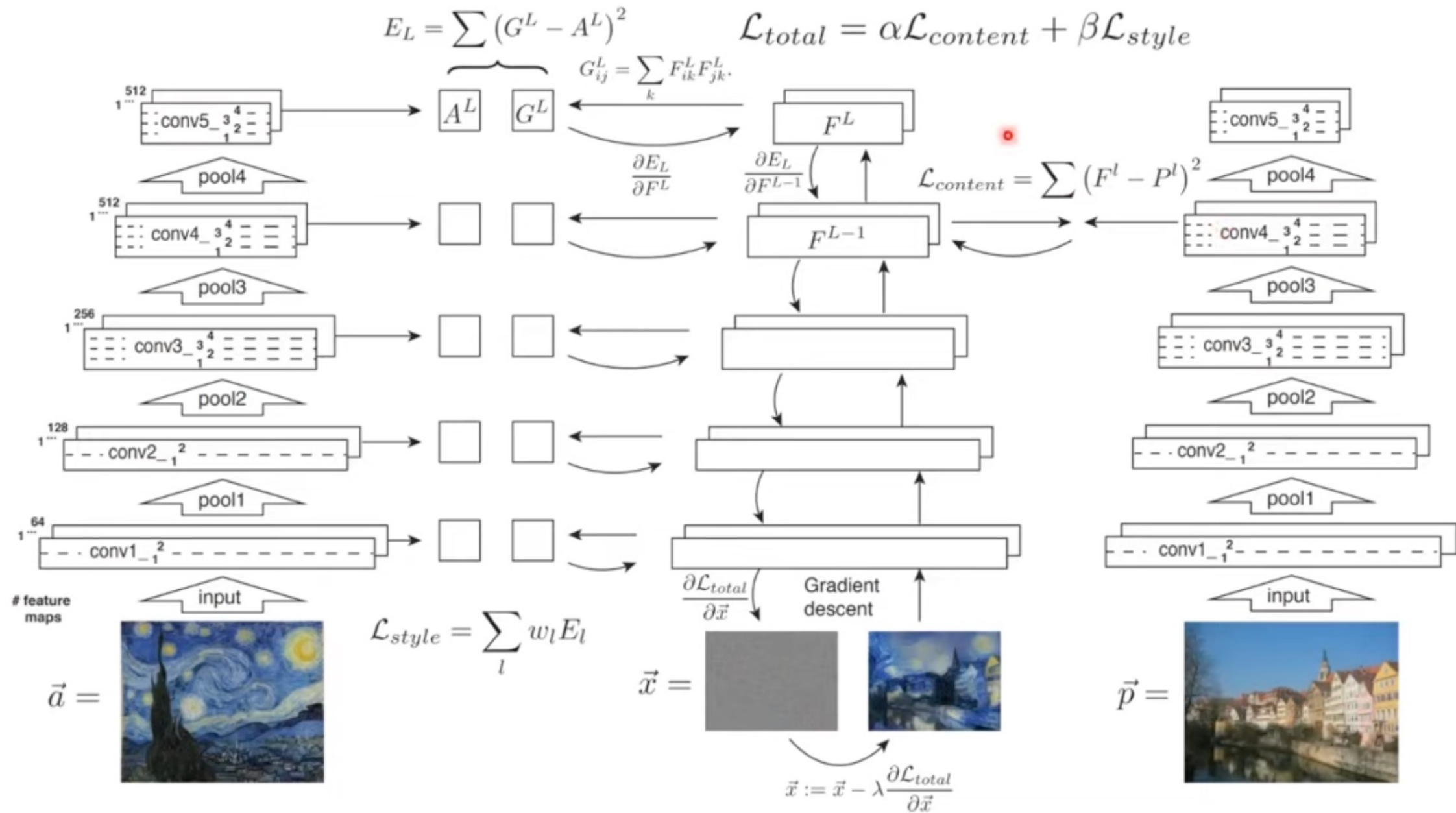
=

Generated Image



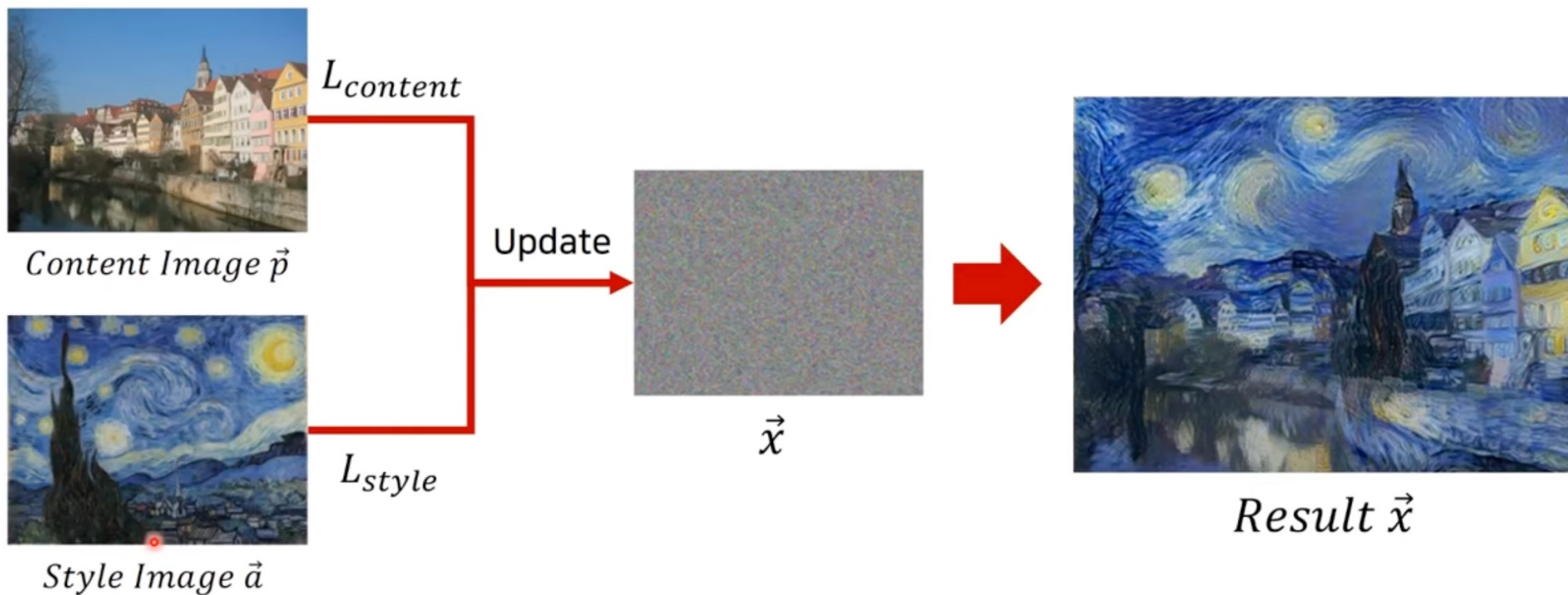
How NST works?

- Neural style transfer uses a pretrained convolution neural network.
- NST defines the following inputs:
 - A content image (c) — the image we want to transfer a style to
 - A style image (s) — the image we want to transfer the style from
 - An input (generated) image (g) — the image that contains the final result (**the only trainable variable**)



Loss Function

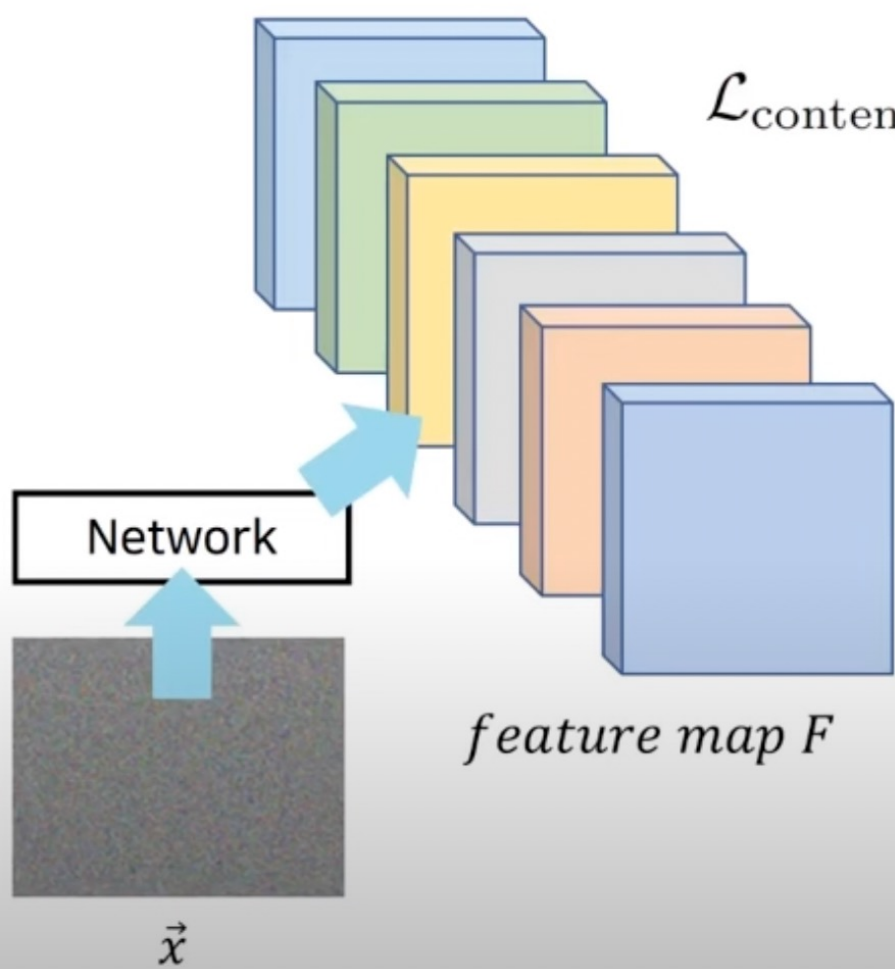
$$L_{total}(\vec{p}, \vec{a}, \vec{x}) = \alpha L_{content}(\vec{p}, \vec{x}) + \beta L_{style}(\vec{a}, \vec{x})$$



What is Content Loss

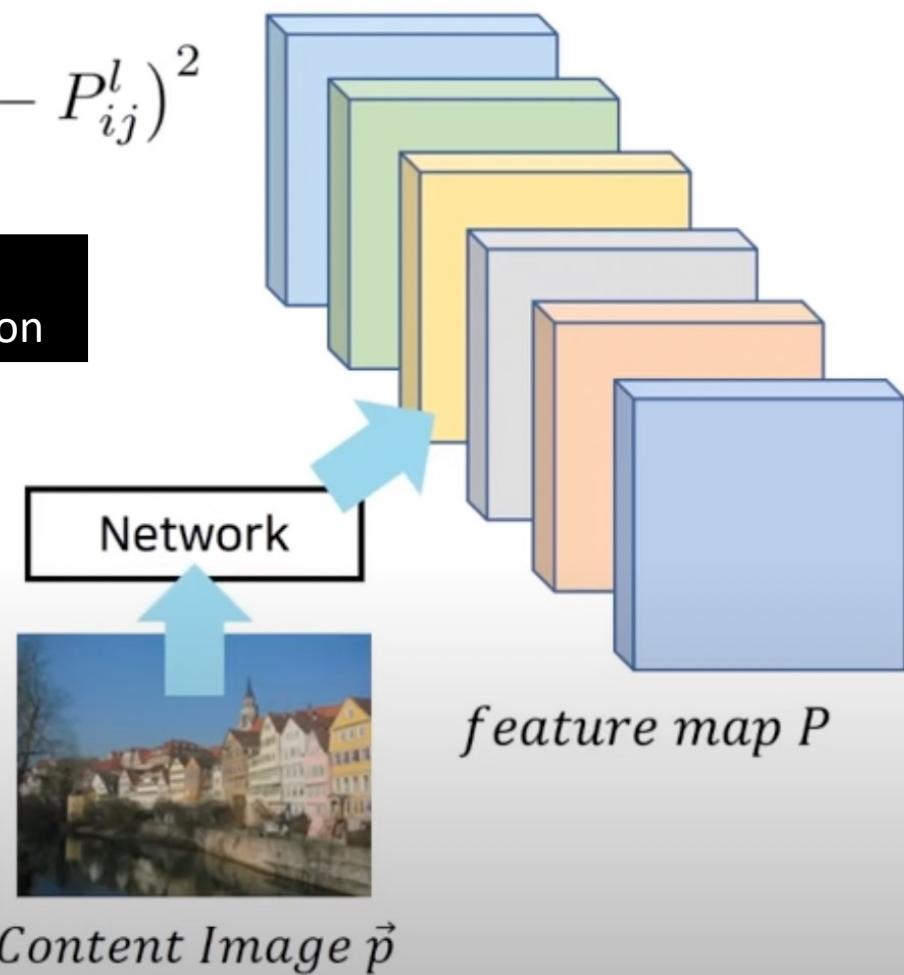
- It helps to establish similarities between the content image and the generated image.
- It has been observed that **CNN** captures information about the content in the higher levels of the network, whereas the lower levels are more focused on the individual pixel values.

$$\mathcal{L}_{\text{content}}(\vec{p}, \vec{x}, l) = \frac{1}{2} \sum_{i,j} (F_{ij}^l - P_{ij}^l)^2$$



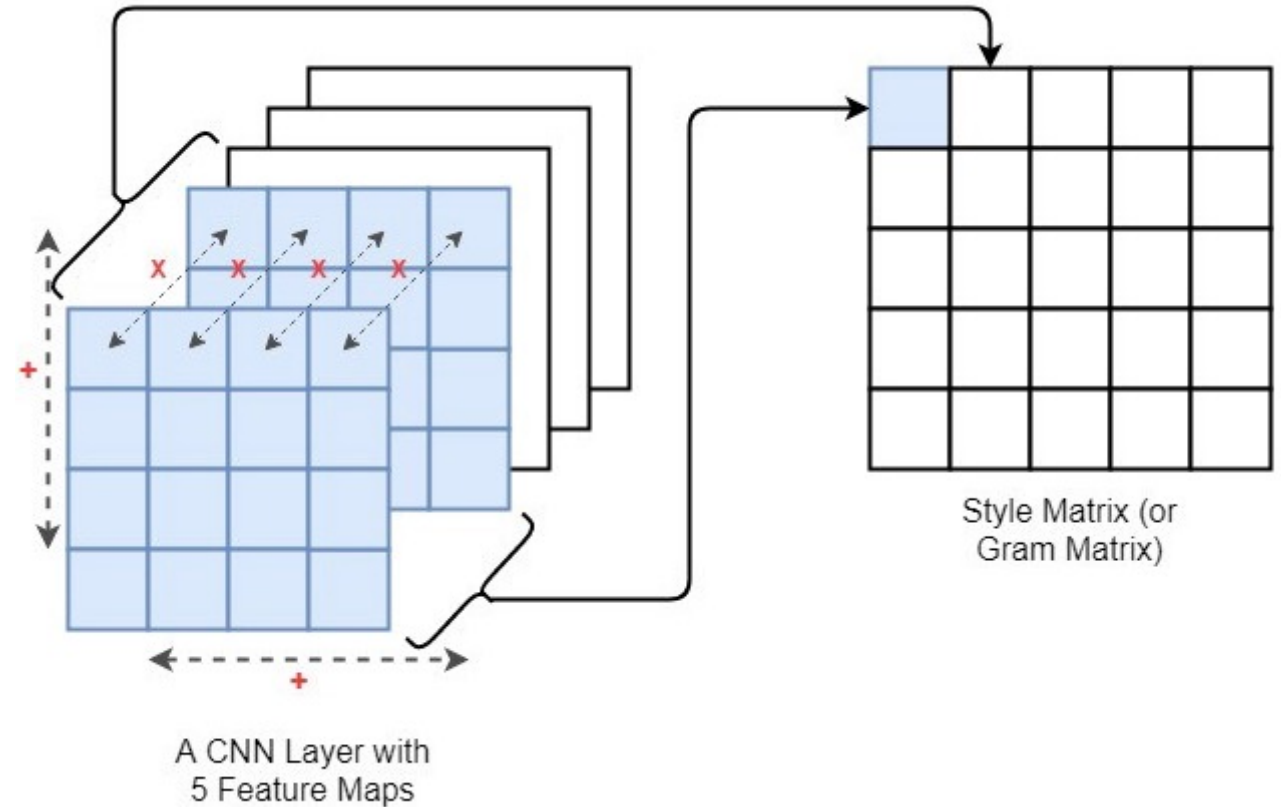
$$\mathcal{L}_{\text{content}}(\vec{p}, \vec{x}, l) = \frac{1}{2} \sum_{\substack{i,j}} (F_{ij}^l - P_{ij}^l)^2$$

i : channel,
 j : location of activation



What is Style Loss

- Gram matrix is a way to interpret style information in an image as it shows the overall distribution of features in a given layer. It is measured as the amount of correlation present between features maps in a given layer.
- Style loss is calculated by the distance between the gram matrices (or, in other terms, style representation) of the generated image and the style reference image.



Summary

- Neural Style transfer builds on the fact to blend the content image to a style reference image such that the content is painted in the specific style
- NST employs a pre-trained Convolutional Neural Network for feature extraction and separation of content and style representations from an image
- NST network has two inputs: Content image and Style image. The content image is recreated as a newly generated image which is the only trainable variable in the neural network
- The architecture of the model performs the training using two loss terms: Content Loss and Style Loss
- Content loss is calculated by measuring the difference between the higher-level intermediate layer feature maps
- Style loss can be measured by the degree of correlation between the responses from different filters at a level.

References

- <https://www.v7labs.com/blog/neural-style-transfer>
- <https://arxiv.org/pdf/1508.06576.pdf>
- <https://www.christies.com/features/A-collaboration-between-two-artists-one-human-one-a-machine-9332-1.aspx>
- <https://towardsdatascience.com/light-on-math-machine-learning-intuitive-guide-to-neural-style-transfer-ef88e46697ee>
- <https://www.analyticsvidhya.com/blog/2020/10/introduction-and-implementation-to-neural-style-transfer-deep-learning/>
- https://www.tensorflow.org/tutorials/generative/style_transfer
- <https://www.youtube.com/watch?v=vaze2c4uKJk&t=123s>